Interaction between Individuals: Its Stability and Style

GORDON PASK
System Research Ltd.
Richmond, Surrey, England

ABSTRACT

Human communication and the stability of interpersonal interaction are considered in a deliberately general fashion in order to highlight features that are common to several types of theory. The concepts in question are unified in a graphical but basically control theoretic notation. Attention is focused on systems that learn and it is argued that if a non trivial learning process is to be represented then the minimal closed and coherent system is, in a certain sense, "Self Reproducing" and that it involves at least two conversationally related individuals. As an alternative to this formulation, it is often convenient, especially in psychological and sociological studies, to insist that the "Self Reproducing" organization is an atomic unit in its own right. This leads us to redefine "individual", as a P-individual (psychological individual). A P-individual is a distinct and manipulable organization but, unlike the usual "individual," it is not generally isolated in a single processor, for example, in a single brain. Rather, a P-individual is a distributed organization and, as such, resembles certain other constructs such as McCulloch's system with "redundancy of potential command" and Vicker's "appreciative system."

1. OVERVIEW OF METHODS AND MEASURING TECHNIQUES

A great deal of psychology and sociology is concerned with measuring and predicting the character of the interaction between individuals. For example, we ask if the interaction is viable; whether the individuals communicate effectively (or are likely to do so); to what extent are they similar beings with ideas in common? Various words are used in the literature to describe the property of consistent and efficient interaction; for instance, "regular" and "harmonious." Probably the best omnibus term to adumbrate this wide range of meaning is *stability* (a *stable* interaction between the individuals concerned) augmented by the auxiliary word *matching* (the individuals are well *matched* if, in certain respects, they are similar).

Stability is a predicate of systems. A system is stable insofar as a salient relationship is preserved during its operation; for example, a vehicle is

Mathematical Biosciences **11** (1971), 59–84

 Copyright © 1971 by American Elsevier Publishing Company, Inc.

stable if certain input/output relations remain invariant. The word may also be used in connection with a system made up from subsystems. The macrosystem is *stable* (literally, the macrosystem *exists*) if and only if the subsystems are related in a particular fashion. For the present purpose, the subsystems are individuals who are potentially or actually interacting with one another. Hence, throughout the article, stability has a broad connotation that will be ostensively defined by the examples in the next section.

After examining the broad implications of stability (and matching), we will look at these conditions and the measuring techniques used to determine their existence in greater detail and, in Section 2, will consider the theoretical commitments of a statement like "this interaction is stable."

A. *Stability of Agreement*

(a) An individual is offered a number of alternatives and is required to choose one of them whenever he is presented with an option. His behavior is stable insofar as his choices satisfy an external standard (in the simplest case, when he makes correct responses). For example, in a classification experiment, the subject's behavior is stable if he correctly assigns items to categories chosen by the experimenter.

(b) The options of (a) become problems and the individual's behavior is stable if he submits correct solutions.

(c) Two individuals, α and β, are posed problems by an external source. Their joint behavior is stable if α and β produce the *same* solutions. In this example, the external standard of (a) and (b) is replaced by a canon of α, β conformity.

(d) α and β are engaged in a conversation about a specified topic. Their interaction is stable insofar as the opinions they express are consistently in agreement. (Very loosely, "the conversation is stable" but this statement is qualified later.)

Several points are noteworthy. First, an assertion of stability is always based upon a comparison of some sort; in (a) and (b) a comparison between an individual's output and a standard, and in (c) and (d) between the output of two individuals. Without exception, so far, the comparison is *behavioral* or *extensional*. This is most clearly manifest in the case of problem solving. α and β in example (c) produce the same solutions, by token of which their mental problem-solving procedures are extensionally equivalent. But they do not necessarily solve the problems in the same *way;* that is, their mental processes are not, in general, *intensionally* equivalent. In example (b) the *standard* is extensionally defined; that is, a solution y is correct if $y = F(x)$ without commenting upon how the function F of the input x is computed. The individual of example (a) is in exactly the

same position, as a result of which it would be illegitimate to infer the existence of a particular concept from the evidence of his stable behavior. Less obviously, perhaps, the α, β agreement of example (d) is extensionally defined; hence, the qualification attached to the usage "this conversation is stable."

B. *Game Stability*

α and β play a game that is explicitly formulated in a normative framework *interpretable* and interpreted by these participants. They both have specified moves; from time to time each participant is impelled to select a move in ignorance of the other participant's action; both of them are aware of the values of a payoff function that is defined over the set of outcomes. The iterated play of the game is stable if α's and β's moves (hence, the resulting outcomes) are invariant or if they belong to a regular and unchanging pattern.

The essential feature of this example is the existence and acceptance of a normative system in which α and β are situated as sentient beings. On closer scrutiny, all the interactions of Section 1A are restricted in the same manner, though the normative constraints are not so well defined. Moreover, though the interactions of Section 1A are frequently gamelike, at least one of them (the conversation of example (d)) corresponds to a very complicated game situation involving cooperation and communication between the players.

C. *Stability of Descriptions and Processes (Intensional Equivalence)*

This type of stability involves not only a matching to determine agreement, but also a comparison of the concepts used in order to reach agreement. The α, β interaction is stable if and only if α and β agree and if they use the same mental processes in order to do so. As a result, the α, β situation must embody an object language that can accommodate statements at various levels of abstraction; in particular, it must be rich enough to accommodate α and β descriptions. It is true that an object language existed in an embryonic form in all of the previous examples (α and β have been making "statements" throughout), but this is the first occasion on which the need for it has become obtrusive.

To illustrate these points, consider a simple concept acquisition experiment. Generally, α is the subject and β is the experimenter (though in a series of group learning experiments due to Lewis and myself [1–3] α and β were members of the group singled out, temporarily, to fill one or the other role). At the outset, a set of objects is specified and a universe of discourse is defined by isolating certain attributes of these objects (indeed, it is usual to generate objects such as Bruner cards by assigning particular

values to each of the attributes). The experimenter β selects a subclass of objects, the target category. Either β presents a sample of objects or α (the subject) is allowed to select one; in either case, for each object examined, β provides α with a signal to indicate that the object in question does or does not belong to the target category. Thus α receives a sequence of evidence on the basis of which he is required to give a description of the target category. His description is stated in terms of an object language interpreted over the universe of discourse and α's description is compared for agreement with β's description of the target category. It should be emphasized that this procedure is quite distinct from the process of β presenting α with an object, asking him to classify it as belonging or not belonging to the target category, and subsequently comparing α's response with (his own) knowledge of membership signal. Of course, α may be requested to perform this trick as well; but object classification statements are at a lower level of abstraction than descriptions.

If α's description and β's description of the target category are (in some sense) the same and if α assigns objects to membership or nonmembership in the same way as β, then (taken over a series of descriptions) the α, β interaction is stable insofar as α and β agree *and* insofar as α and β have the same grounds for agreement. Moreover, object classification is a form of problem solving, since any description expressible in the matrix form of Hunt *et al.* [4] may be rewritten as the test tree for a problem-solving procedure and any relational description is subject to similar treatment. Hence, the dictum about α and β can be rephrased, "α and β submit the same solutions *and* solve the problems in the same way." But these comments gloss over the important fact that there are several cogent methods of comparing descriptions and, consequently, several senses in which descriptions may be "the same."

To see this, notice that even though the universe of discourse is circumscribed, several *types* of description exist; for example the type of description encompassed by the first-order predicate calculus of logic (the target category is described by a statement involving the values of unary predicates, the attributes, together with the usual logical connectives) or the relational type. It is customary also to distinguish *subtypes*; for instance, "conjunctive descriptions" (involving only the connective "and"), disjuncts of conjunctions, implicative descriptions, and so on. Of course, instructions may be given or the object language may be restricted so that only one subtype (e.g., "conjunctive") is permitted. But even in this case, there is generally a many-to-one relation between the possible, sample-based α descriptions of the target category and β's description of it.

Hence, the comparison of α and β descriptions depends upon what *degree* of *intensional* equivalence constitutes a *match*. Is it sufficient if α and

β produce descriptions, unqualified, of the same target category; must the descriptions be of the same type; must they have the same subtype; must they be identical? To define a stability criterion, it is necessary to answer these questions. In general, the wider α's and β's interpretative capabilities, the more liberal the answer will be.

D. *Conceptual Matching*

Sometimes α, β interaction is deemed stable (roughly, in the sense that α and β understand one another) if α and β have the same sorts of concept, regardless of whether or not α and β are in accord. In a rather narrow sense, the existence of this type of α, β matching determines the possibility of α, β communication with respect to a given universe of discourse.

To secure an index of matching, I use an interrogation method that is a variant of the Kelly grid technique [5, 6]. First of all, a universe of objects is selected. These are quite arbitrary (people, snowflakes, the leaves of trees, objects pertinent to any topic that α and β may talk about). As in administering the standard Kelly grid, triads of objects are presented (separately) to α and β. Each individual states an attribute in respect to which two members of the triad are similar and one is different. Each of these attributes is entered as a row name in an array that has columns labeled by the objects A, B, ... The cells of a row are filled in (for α) by asking him to state (for each object) that it has the attribute $(+1)$, that it does not (-1), that the value is undetermined (0), or that the attribute is meaningless or inapplicable (*). An array with rows $\alpha_1\ \alpha_2 \cdots$ is generated by repeatedly questioning α and a similar array with rows $\beta a, \beta b, \ldots$ is generated for β; for example, the arrays

A	B	C	D	E			A	B	C	D	E	
[+1]	[−1]	[+1]	*	−1	α1	and	[+1]	[−1]	[+1]	−1	−1	βa
[−1]	[+1]	*	[+1]	0	α2		[−1]	[+1]	*	[+1]	−1	βb

Clearly the triads (enclosed in □) must have two entries marked with a $+1$ and one with a -1. The α attributes are now presented to β for judgment and vice versa, thus providing two tables, namely, an α array and a β array, such as

α array

A	B	C	D	E	
[+1]	[−1]	[+1]	*	−1	α1
[−1]	[+1]	*	[+1]	0	α2
+1	+1	0	0	−1	αa
+1	−1	0	*	*	αb

β array

A	B	C	D	E	
[+1]	[−1]	[+1]	−1	−1	βa
[−1]	[+1]	*	[+1]	−1	βb
+1	−1	−1	*	−1	$\beta 1$
−1	+1	*	0	0	$\beta 2$

A crude index of α, β matching is obtained by rank correlating the non * (meaningless) entries in these arrays and by ascertaining the number and distribution of the various * entries.

The outstanding feature of this example is that there is no standard "correct description" (in the sense that a target category description is correct and arrived at by the proper interpretation of evidence). Nor are the relevant attributes determined a priori. However, the individuals do end up by stating a description (which, like the matrix form of a concept description) can be refined to specify a mental concept recognition procedure and they do so by making statements in an object language that is inherent in the interrogation method. With a slight extension of this method (Hinkle's [7] laddering technique) α and β can be persuaded to make more abstract statements in reply to "why" questions, for example, "Why do you assign the values $+1$, -1, $+1$, *, -1, to $\alpha 1$?"

E. *Interpersonal Stability*

Conceptual matching says something about α's understanding of β and β's of α but it does not fully characterize effective communication between α and β. To do so, it is necessary to ask a number of questions, perhaps an indefinitely large number. For example, do α and β agree? Does α understand β? Does β understand α? Do they appreciate their understanding? For instance, it is entirely possible for α and β to communicate in the absence of immediate understanding if, on the one hand, α appreciates that he misunderstands β and vice versa. But if misunderstanding exists, and it is not mutually appreciated, then communication is effectively blocked.

The problem of determining what appreciations and understandings exist has been tackled in the clinical context, notably by psychologists such as Bateson [8], Brodey [9], and Laing [10], who are interested in familial schizophrenia and kindred phenomena. By far the most elegant test instrument is the interpersonal method (IPM), due to Laing *et al.* [10]. In its original form, the IPM is concerned with family relations (α husband, β wife). The test is administered to α and β individually. It posits a sequence of 60 issues that are deeply bound up with the α, β relationship; for example, the issue of "depending upon." For each issue, α replies to enquiries of the form: How true are the following statements?

"β depends on me (α)";
"I (α) depend on β";
"β depends on herself (β)";
"I (α) depend on myself (α)";

a rephrased series being presented to β. The IPM is also designed to elicit

metastatements: thus, α is asked how α thinks β would reply to an appropriately pronominalized version of the original enquiries. Finally, the IPM elicits meta-metastatements by asking α "How do you think β thinks you would reply to" the same question set.

The analysis of IPM data is quite complicated, since all of the clinical issues involve polar α, β relationships at the lowest level. But very much the same technique can be employed when it is desired to elicit α's and β's opinion about some object or situation that they are both able to observe and I have used a (modified) IPM quite extensively in this fashion [11, 12]. Typically, the issue Z is of the form "Which event will take place?" (naming the events) or "Which alternative?" naming the moves in a game-like situation.

Each individual responds to a series of questions of this type, one for each value of Z. Let us denote the responses $\alpha(Z)$ and $\beta(Z)$, depending upon whether they are obtained from α or β. Thus:

$\alpha(Z)$: α's response to "α! What do you think about Z?"

$\beta(Z)$: β's response to "β! What do you think about Z?"

Each participant also makes a series of metaresponses, namely,

$\alpha[\beta(Z)]$: α's response to "α! What do you think β thinks about Z?"

$\beta[\alpha(Z)]$: β's response to "β! What do you think α thinks about Z?"

Finally, both α and β respond to meta-metaenquiries:

$\alpha[\beta[\alpha(Z)]]$: α's response to "α! What do you think β thinks you (α) think about Z?"

$\beta[\alpha[\beta(Z)]]$: β's response to "β! What do you think α thinks you (β) think about Z?"

These responses are quantified by enforcing a selection among the alternative moves, by obtaining a (subjective) probability estimate for each one or by a combined sampling method (introduced in Section E). Hence, α's responses may be compared with β's and vice versa. Following Laing, with some small and contextually justified deviations, we define:

1. α, β agreement = comparison of $\alpha(Z)$; $\beta(Z)$;
2. α's understanding of β = comparison of $\alpha[\beta(Z)]$; $\beta(Z)$;
3. β's understanding of α = comparison of $\beta[\alpha(Z)]$; $\alpha(Z)$;
4. α, β understanding = [α's understanding of β] $\wedge$ [β's understanding of α];
5. α's appreciation of β = comparison of $\alpha[\beta[\alpha(Z)]]$ $\beta[\alpha(Z)]$;
6. β's appreciation of α = comparison of $\beta[\alpha[\beta(Z)]]$ $\alpha[\beta(Z)]$;
7. α, β appreciation = [α's appreciation of β] $\wedge$ [β's appreciation of α];

8. α's impression of agreement with β = comparison of $\alpha[\beta[\alpha(Z)]]$; $\alpha(Z)$;

9. β's impression of agreement with α = comparison of $\beta[\alpha[\beta(Z)]]$; $\beta(Z)$;

10. α's impression of understanding β = comparison of $\alpha[\beta[\alpha(Z)]]$; $\beta[\alpha(Z)]$;

11. β's impression of understanding α = comparison of $\beta[\alpha[\beta(Z)]]$; $\alpha[\beta(Z)]$.

In the game-like situation, for instance, a "comparison" is interpreted as joint (subjective) probability distribution over the outcomes generated by the possible moves of α and β.

It is contended that the stability of an α, β conversation generally depends upon all of these comparisons and not just upon agreement or understanding alone. The cogency of this remark will become evident in Section 2. For the moment, let us focus our attention upon a crucial feature of the IPM itself; namely, that the responses are *metalinguistic*. Thus, α is asked to make statements *about* β's statements and his reasons for making them and indirectly he is asked to comment on his own mentation.

To see this, notice that a statement of the form

$$\alpha[\alpha(Z)] \quad \text{or} \quad \beta[\beta(Z)]$$
$$\alpha[\alpha[\alpha(Z)]] \quad \text{or} \quad \beta[\beta[\beta(Z)]]$$

is an introspection in the classical sense of a self-observation. Logically, this is the simplest sort of metastatement, though psychologically speaking, it is notoriously difficult to elicit. Indeed, when subjects are trained to act as observers, they become able to isolate essentially two individuals embodied in the same brain; I now (the observer) and I there (the person who is observed). The separation of two or more individuals, one who speaks and one who is spoken about, is at the core of metalinguistic utterance.

These metallinguistic statements, it should be emphasized, are quite different from the object language statements elicited by most of the techniques described previously. In particular, the levels of the IPM (which will later be identified as levels of *control*) are quite different from levels of *abstraction* (i.e., the level of descriptive statements as against the level of object statements, in Section 1C). There is only one exception: the "why" questions of the laddering technique (which yield responses at a level of abstraction above the Kelly grid description) are also introspective metastatements. In responding "because . . .", the subject (now) is making a statement about the process that was himself a moment ago (when he filled in the grid).

F. *Combined Sampling*

IPM response format may literally *be* a grid (used in the reentrant form outlined in Section 1D). With this combined sampling method, hierarchies of control and of abstraction, though quite distinct, may be jointly investigated. α is asked to fill in a grid personally, responding with $\alpha(G)$; to fill it in as he believes β would fill it in, responding with $\alpha[\beta(G)]$; to fill it in as he believes β thinks he would fill it in, responding with $\alpha[\beta[\alpha(G)]]$. A similar set of grids is obtained from β and IPM comparisons are made in terms of G entries rather than Z scores. In addition, the laddering method can be employed (in practice, easily enough) on $\alpha(G)$ and $\beta(G)$ to yield introspections $\alpha[\alpha(G)]$ and $\beta[\beta(G)]$, which are usefully compared with $\alpha[\beta(G)]$ and $\beta[\alpha(G)]$. The comparative measures obtained from combined sampling are tolerable predictors of effective communication between α and β.

Discussion

We have now exhibited five separate uses of the terms "matching" and "stability" and have briefly reviewed the pertinent measuring techniques. Many practically important human systems involve all these types of stability. One example, especially relevant to the discussion in Section 2, is a system for computer-assisted instruction.

The five sorts of stability enter the system designer's mind in the following ways.

(a) The student (α) must be *marked* as right or wrong when he solves problems that are presented by the teaching device. *Marking* involves the comparator of Section 1A(b). Knowledge of results is provided by the extensional comparator of Section 1A(c), with the teaching device playing the role of β.

(b) Nontrivial teaching systems are gamelike insofar as they are set up in the normative scheme of a tutorial contract with the student, thus invoking the game stability of Section 1B.

(c) If the teaching system stabilizes the student–teacher interaction by providing differential cueing or problem simplification, it (β) must contain a model for *how* a student (α) solves problems. Otherwise it could not offer an individually appropriate form of cooperative assistance (cueing, simplification, or whatever) and unguided efforts in this direction would be just as likely to confuse the student or even to mislead him. The apposite matching criterion is the intensional equivalence of Section 1C and the tutorial dialogue generally takes place at several levels of abstraction.

(d) Effective representations of subject matter, knowledge, or skill structure, must be conceptually matched to the student's mind, in the sense of Section 1D.

(e) "Conversational" teaching systems involve metalinguistic discourse concerned with the direction of attention and the selection of goals. The student (α)–teacher (β) matching process [12] involves a sampling arrangement that is isomorphic with the "combined sampling" of Section 1F.

Concern over diverse forms of stability and matching is thus an entirely practical concomitant of tutorial design and it seems likely that similar considerations will have just as much pragmatic consequence in areas such as information retrieval, where human beings interact with computer-instrumented storage facilities.

It is no accident that most of the ideas and methods had their origin in the social sciences, where the demand for sophisticated concepts of stability is obvious and prodigious. Unfortunately, the systems encountered in that domain are, by their nature, logically and mathematically intractable. As a result, the value of the ideas and methods that did emerge has often been obscured. They have been dubbed "qualitative not quantitative" or sometimes "mentalistic." (I am prone to regard the last property as a virtue, but the word is often used in a pejorative manner.) Nowadays, these same notions are being applied to quasi-mechanical systems (as in the last paragraphs; teaching systems, information retrieval systems) where, in the past, the human constituent has been treated in an absurdly simpleminded fashion. In such fields, logical and mathematical models, albeit rather untidy ones, *can* be erected. These, in turn, may buttress the sophisticated concepts and render them formally acceptable. The theoretical discussion in the next part of the article takes a small step in this direction.

2. THEORIES

The following subsections sketch a number of psychological or sociological theories and point out that they differ in respect to the notions of stability that they can accommodate. It also turns out that their power in this respect is intimately related to their capacity for representing nontrivial modes of learning. Consideration of this topic leads, in Sections 2G and 2H, to a reappraisal of the usual concept of an individual.

A. *Strict Behavioral Theories*

An individual is conceived as a black box, in the sense of Ashby [13]. An experimenter observes the behavior of one or more black boxes (labeled α and β) and infers their transfer functions. One type of observation consists in a comparison of the conjoint output states of black boxes α and β, subject to a common forcing input V. The output of the comparator is U, often the equality function or agreement index of Section 1A,

and the entire procedure is shown, using a standard notation, in Fig. 1. The macrosystem α, β is stable if and only if U assumes specified values. This is the *only* type of stability that is accommodated by a strict behavioral theory. Moreover, it should be noted that this canon of stability (in common with the identification of individuals, their input and output variables, etc.) stems from the experimenter and is defined in *his* descriptive metalanguage in which neither α nor β is allowed to speak. The experimenter justifies his choice of criteria, variables, and so on in terms of observations that are substantially independent of the interaction depicted in Fig. 1; for example, in terms of the physical character of α and β, in terms of drive reduction, or the like.

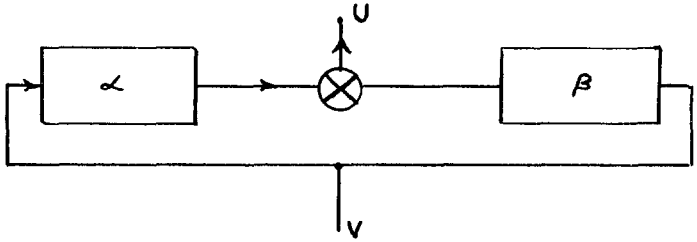

FIG. 1. Black box interaction.

B. *Behavioral Theories of Goal-Directed Systems*

Within the compass of a strict behavioral theory, it may be legitimate to say that α and β behave as *though* they aim for certain goals; it would not be legitimate to say that they *are* goal-directed systems. Perhaps the

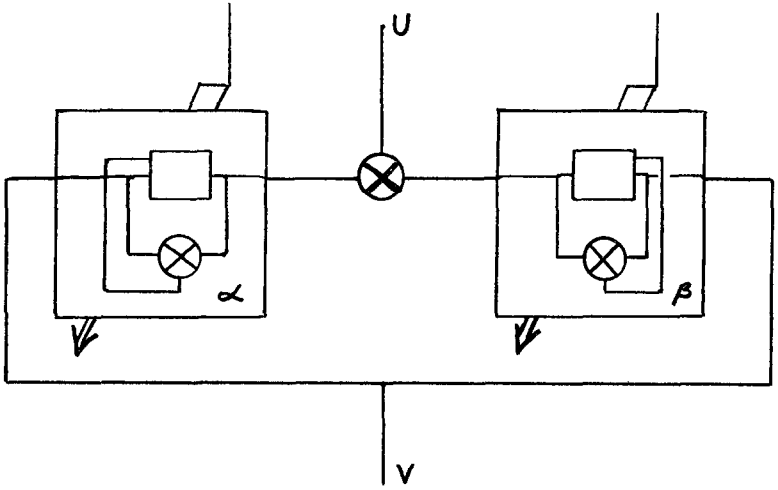

FIG. 2. Interaction of goal-directed systems.

weakest grounds upon which a statement of goal directedness can be based are as follows: α and β are *prescribed* or *constructed* so that they maintain certain input/output relationships (they are automata built for that purpose) or (alternatively) the experimenter has examined the mechanism inside the black boxes and has *discovered* a structure that *is* such an automaton. Given a structural theory, the form of the comparison in Fig. 1 can be justified without resorting to independent evidence. Equipped with structural knowledge, the experimenter can also distinguish certain variables as parameters, as suggested by Fig. 2.

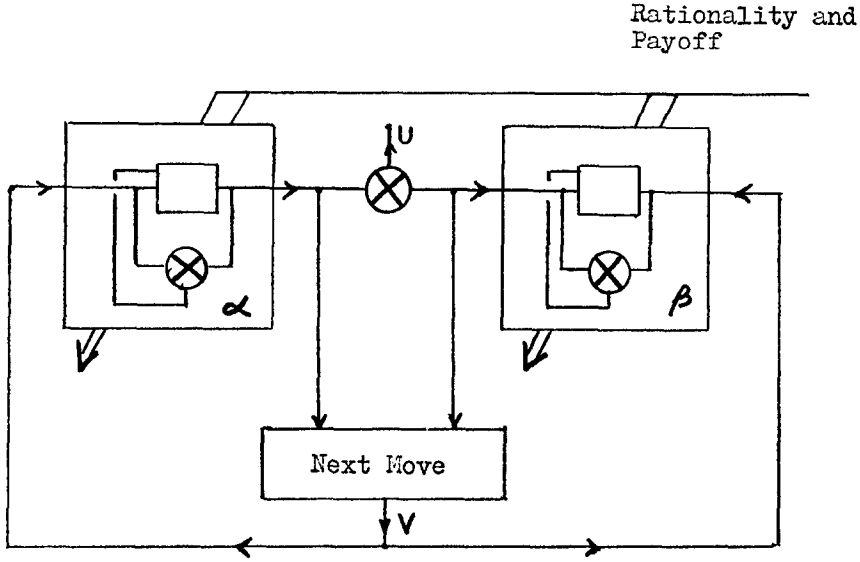

FIG. 3. Normative interpretation. The output of the internal comparator in each unit is an abstractive statement of the input/output relation used to indicate goal achievement.

C. *Normatively Interpreted Theories of Goal-Directed Systems*

If α and β are regarded as creatures who are able to understand and interpret a language that is also spoken by the experimenter, then the parametric input of Fig. 2 may represent an *instruction* to aim for a certain goal or it may represent an *agreement* among α, β, and the experimenter to abide by a contract. In these circumstances, it makes sense to talk about α and β as game players and to interpret U as the outcome function comparison of Section 1B. In this case, the parametric input also stipulates the payoff function of which α and β have full knowledge (Fig. 3).

Such a normative interpretation of the system is the license needed in order to postulate a principle of rationality in the usual game-theoretic

sense (or to impose it as a principle α and β *should* adhere to, which is a further parametric constraint). If α and β are rational beings, then game theory may be used to *predict* the occurrence of stable behaviors; in general, these correspond to the Nash equilibria [14].

D. *Extension of the Normative Theory*

The contract of Section 2C entails a goal (to maximize payoff while acting prudently; to be a rational player in full knowledge of the payoff function). It also gives the status of an object language to the channel of α, β interaction.

In the context of game playing, this remark is trivial. The object language statements emitted by α and β in the course of a simple game consist in move selections. The remark is nontrivial insofar as the channel conveys messages at different levels of abstraction that are compared by different mechanisms. Figure 4 extends the notation to incorporate this refinement and thus to encompass the types of stability mooted in Section 1C. Figure 4 is equally well interpretable as a game situation or a problem situation (Bannerji [15] establishes the requisite formal equivalence between game playing and problem solving.)

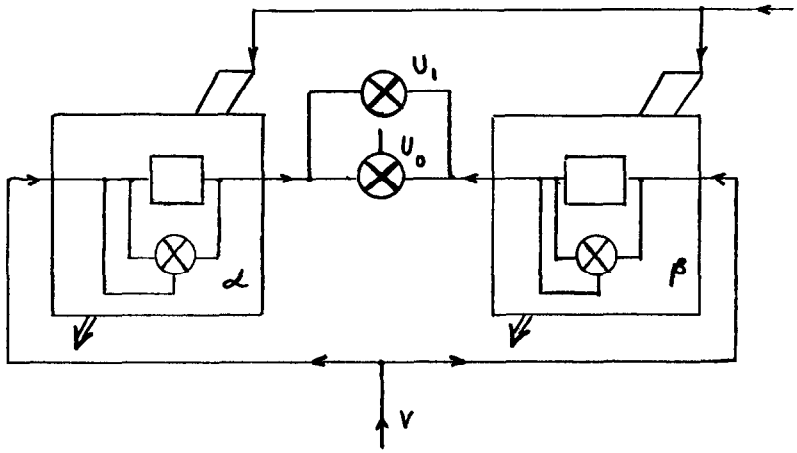

FIG. 4. The status of two levels of abstraction with matchings U_0 and U_1. Notice that U_0 and U_1 are at different levels of abstraction and should not be confused with the superscript notation for levels of control (i.e., U^0, U^1) that is employed in the other figures.

E. *Theories Involving Hierarchies of Control- or Goal-Directed Organization*

Single-level theories are not rich enough to represent the realities of choice and metacommunication or to give a proper picture of processes

such as attention directing or learning. A number of multilevel theories have been produced in response to the challenge provoked by these deficiencies. These theories are of two closely related types: *formal* theories and *operational* (or psychological) theories.

i. *Metagame Theory*

The most incisive formal development is due to Howard [16, 17] and is set in the game-playing context of Sections 1B and 2C. As a preliminary to his discussion, Howard notes that certain games are paradoxical: Prisoner's dilemma, with the payoff matrix shown here, is a classic case (only the ordering of the entries being important).

		β X	β Y
α	X	3, 3	1, 4
	Y	4, 1	2, 2

The paradox is due to the fact that the Nash solution (the rational solution) is counterintuitive. It yields $\langle 2, 2\rangle$ whereas the players ought, in a perfectly obvious sense, to achieve $\langle 3, 3\rangle$. Further, in iterated plays of the game, some players *do* achieve $\langle 3, 3\rangle$ as a stable mode of operation (some others, it is true, reach stability at $\langle 2, 2\rangle$) and a competent theory should predict both sorts of stability.

The situation is resolved by noting that even the single-level prediction of the Nash solution rests upon the assumption that α and β entertain a particular hypothesis (Howard calls this a "theory"; I have substituted "hypothesis" to avoid confusion due to the wording of this article) that is, of first-level rationality. Thus α chooses move Y because he believes that β, using the same hypothesis, will also choose Y. True, in a single-level theory, this hypothesis is excluded from explicit consideration by the gambit of writing it in the form of a parametric input (Fig. 3), which is *displaced* from the theory itself. But that is beside the point.

In Howard's theory α's hypothesis is represented as a mapping from β's moves to α's moves. Now there are four such mappings, which may be summarized as (a) = "Choose X always"; (b) = "Choose Y always"; (c) = "Choose the same as β"; (d) = "Choose the opposite of β"; and it is presumed by an external observer that player α may *choose between* these hypotheses just as well as he may opt to *choose one* of them. Operationally, this implies (1) that α chooses among a set of conditional strategies (leading to the selection of X or Y) rather than choosing between X and Y per se, and (2) that α does so on the basis of a *metamodel* for β's choice

(i.e., a model for the class of models β might employ). Formally, the construction yields a 2×4 outcome matrix characterizing a *metagame.*

With due reservations, Howard imports an assumption of payoff invariance; that the payoff values remain the same regardless of how or why the move was selected. Given payoff invariance, the cells of the metagame matrix are filled as follows.

	X	Y
a	3, 3	4, 1
b	1, 4	2, 2
c	3, 3	2, 2
d	1, 4	4, 4

There is still only one Nash equilibrium. However, the construction may be repeated by supposing that β is able to choose a way of reacting to α's way of reacting to β's actions. Operationally, this implies (1*) that β chooses between a set of 16 conditional strategies (all of which lead to a selection of X or Y) rather than choosing between X and Y per se, and (2*) that β's choice is based upon a metamodel of α's metamodel of him. At this point a formal representation yields a 4×16 outcome meta-metagame matrix, the cells of which may be filled, adhering to payoff invariance, with the original payoff pairs. Now, in this matrix, there are three Nash equilibria, two of them yielding the intuitively maximal payoff of $\langle 3, 3 \rangle$ and one of them yielding a payoff of $\langle 2, 2 \rangle$. The construction can be repeated indefinitely but Howard shows that with two players and the assumption of payoff invariance, no further solutions will be uncovered. (In general, for n players, the $n!$ expansion is exhaustive.)

Howard's theory of metagames is quite generally applicable. It may be applied in the case of n participants, $n > 2$, or to choice by a single individual. For zero-sum games, it has little interest, since it leaves the value of the game unchanged; "value" being used in its technical sense. For partially cooperative games, on the other hand, it usually does reveal novel solutions. It is a very subtle theory, no more than glossed in these paragraphs; the reader is invited to peruse the original papers [16, 17] and a very perspicacious article by Rappaport [13].

ii. *Implications*

The theory has formal implications (structuring of metagames and predicting their equilibria) and psychological or operational implications to do with the way that metagames are played. Formally, the theorist is

invited to consider strategies that are rational in the light of classes of hypotheses entertained by α and β (i.e., in the light of models, metamodels, and meta metamodels). Operationally, the theory implies that α and β base their game-controlling actions on the hierarchy of models. In fact, the theory exhibits this scheme, as an operational (psychological) *possibility;* it says nothing directly about the *mechanisms* involved. But an external observer, in the position of a formal theorist, is certainly led to make certain measurements that refer to these mechanisms and thereby to predict what stability will *actually occur* in the α, β interaction (not merely the *equilibria that exist*). That is, he may select among the rather large number of possible equilibria that are *Howard* predictable. Moreover, the existence of a comprehensive formal theory invites us to construct operational or psychological theories that model the actual process of game playing.

iii. *Measurement*

The requisite measuring technique was introduced in Section 1E. It is the IPM scheme and is control-theoretically diagrammed in Fig. 5. The otherwise arbitrary truncation of the IPM procedure at the meta-metalevel is now explicable as a tacit assumption of payoff invariance (or, in situations that are not gamelike, of value invariance), which, on consideration, *does* go along with most applications of the method.

iv. *Operational Models*

Any model compatible with metagame theory must credit the individual with an ability to choose hypotheses; either to construct hypotheses *de novo* or to select one from a set of them for immediate currency. In psychological terms, these operations are either *learning* operations or *attention*-directing operations (the distinction is relatively unimportant; it depends upon how an individual is observed and the storage facilities he is supposed to possess). Hence, multilevel theories are potentially theories of learning or attending, in which the structural skeleton of an individual is usually represented by the *control* hierarchy of Fig. 6.

Each level of control $L^0, L^1, \ldots$ corresponds to a *metalevel* in a language that might be used for communication between several individuals, $\alpha, \beta, \ldots$ and the boxes themselves contain models or metamodels. For example, the input to the L^1 box in Fig. 6 is a descriptive metastatement *about* some individual (or model for an individual). The output from the L^1 box has the logical stature of a *command* in Rescher's [29] sense (that which sets up a process in an individual; that which is issued to someone upon someone's *authority* and which may (or may not)

be terminated, that which is justified by a *reason* for giving the command; an explanation) a command is addressed to some individual and, if internally directed, is a *constructive* operation. Depending upon the interpretation of Fig. 6, this constructive operation may either be an *attention-directing* operation or an imperative (model-changing) operation involved

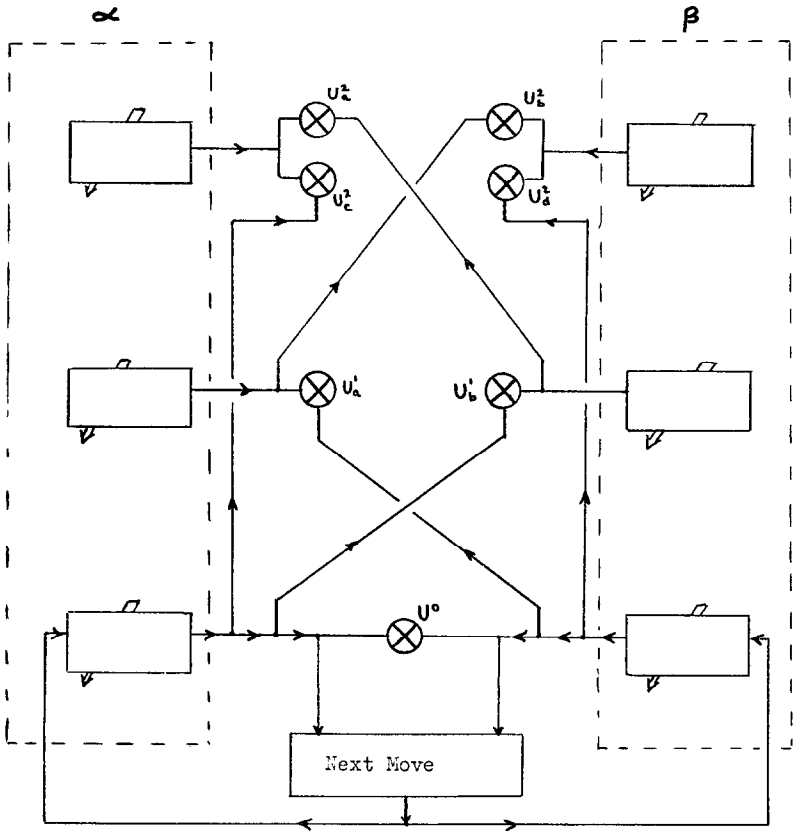

FIG. 5. A control-theoretic representation of the IPM measures of Section 1C. The boxes α and β have the contents of Figs. 2 and 3 but the detail is suppressed in this and all subsequent figures. Notation:
$U^0 = \alpha(Z)\colon \beta(Z)$, $U_a^1 = \alpha[\beta(Z)]\colon \beta(Z)$;
$U_b^1 = \beta[\alpha(Z)]\colon \alpha(Z)$, $U_a^2 = \alpha[\beta[\alpha(Z)]]\colon \beta[\alpha(Z)]$;
$U_b^2 = \beta[\alpha[\beta(Z)]]\colon \alpha[\beta(Z)]$, $U_c^2 = \alpha[\beta[\alpha(Z)]]\colon \alpha(Z)$, $U_d^2 = \beta[\alpha[\beta(Z)]]\colon \beta(Z)$.

in nontrivial learning. It should be emphasized that levels of control and metalevels in Fig. 6 are levels of choice, choice about choice (as in a metagame) but they are not identical with the levels of abstraction in Fig. 4 and

do not generally correspond to them. To go up one level of control, an individual must speak about or operate upon another individual. To abstract, he need only form a class of things.

These preliminary comments have already uncovered a curious circularity in the concept of an "individual," as a result of which the comparators of Fig. 6 are ambiguous symbols. The ambiguity will not

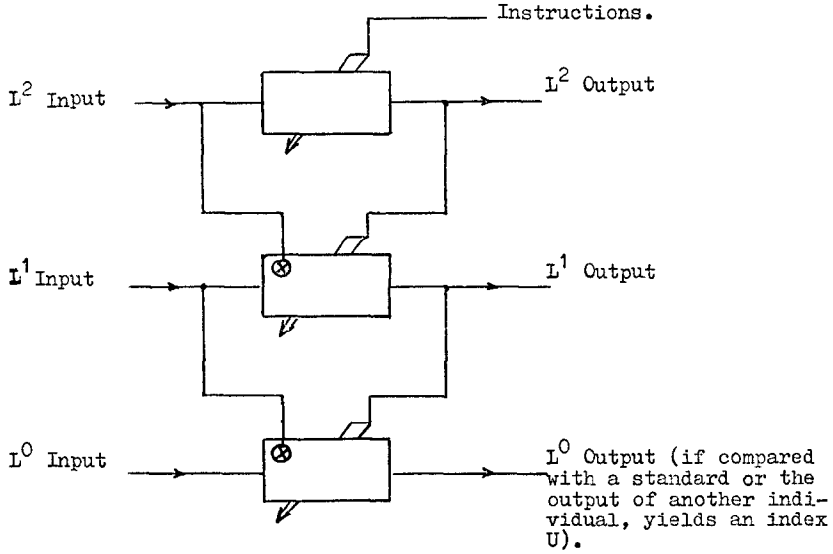
Instructions.
L^2 Input
L^2 Output
L^1 Input
L^1 Output
L^0 Input
L^0 Output (if compared with a standard or the output of another individual, yields an index U).

FIG. 6. Outline of a Multilevel model.

be resolved at this juncture; we will return to the matter somewhat later when better equipped to deal with it. For the moment, the comparator notation is roughly interpreted as follows. The output of a comparator is a descriptive metastatement about the box in which the comparator resides. A complete loop spanning two levels (i.e., a metastatement from a comparator, a higher-level box, and a constructive operation acting on a lower box) constitutes an *explanation*. Further, this is the least organization that accommodates an *explanation*. We say that the boxes represent models and metamodels just because explanatory loops surround each one of them.

A number of cognitive theories fit easily into the framework of Fig. 6, notably Kelly's theory of personal constructs [6] and Festinger's theory of cognitive dissonance [19]. So far as I know, only two of these theories, von Foerster's [20, 21] and my own [12, 22, 23], make the relation to metagames and metatheories fully explicit. Both of these theories are presented as theories of "learning *and* teaching," though both of them can also be

interpreted as theories of *conversation* between *two* participants α and β. This dualism or relativism is not accidental. It is necessary to invoke α *and* β or teacher *and* student in formulating the pertinent ideas.

(In this connection it should be emphasized that learning and teaching are not conceived as a species of adaptation, conditioning, or the like. They are supposed to consist in a hierarchy of problem solving or, since the canonical images of problem solving and control are isomorphic, in a hierarchy of control processes. Storage and retention mechanisms exist, but their nature is not explicitly considered. The theories concentrate on the *form* of learning and on organizational properties, such as *memory*.)

v. *Von Foerster's Theory and My Own*

Von Foerster takes finite-state machines (FSMs) as his building blocks. These exist at the lowest level in the system, L^0, as problem solvers or controllers (i.e., a structural interpretation is available in the sense of Section 2B, so that an observer can say "This automaton is goal directed"). At the next level, L^1, there are also finite-state machines (FSMs) with the same interpretation, but these act upon a domain of L^0 finite-state machines and modify the next state functions of the L^0 FSMs, literally constructing new states by modifying old ones. Because of this fact, any of the L^1 FSMs is called a finite function machine. Von Foerster has developed the mathematics of the model and has interpreted it chiefly in the context of neuro physiology.

As a slight generalization, let the FSM be an arbitrary automaton, sited at level L^0, in which case the simplest finite function machine is an L^1 automaton that *reproduces* this L^0 automaton in the sense of von Neuman [24]. Hence, in Fig. 6 a complete loop spanning levels L^0 and L^1, starting at a comparator and ending at a constructive operation, is a *reproductive* process. It is also, as previously pointed out, an *explanation* (Loefgren has recently highlighted a formal equivalence between "reproduction" and "explanation").

The *operation* of the L^1 automaton, when it is engaged in veridical reproduction, is defined as the individual's *memory* of whatever L^0 automaton is reproduced (the distinction of "memory" as an *organization* is crucial; memory is *not* "storage"). For example, the L^0 automaton may be the *process* called a *concept* in Section 1, when the L^1 operation is the memory of this concept.

If the reproductive cycle is perturbed by variation and constrained by principles of cooperation and selection, the reproductive process becomes evolutionary and its activity is manifest as *learning*. This learning is defined as a form of evolution, not of adaptation.

One embodiment of von Foerster's theory is due to Brieske [25].

Another one is a programme by Fogel, Owens and Walsh [26]; FSMs addressed to predictive problem solving in a symbolic environment evolve under the control of an executive that serves as a finite function machine. A few other programs, having a similar caliber, are reviewed elsewhere [27] and very similar concepts were expounded by Newell and his colleagues several years ago [28].

My own theory is less precise, but more eclectic, insofar as it starts with the generalization noted in the last paragraph. The building blocks (GDSs, goal-directed systems) are procedures, either serial or concurrent, interpretable by a general-purpose computing machine with a suitable programming language (an object language in the sense of Section 1). Among other physical systems, parts of the brain are supposed to act as the machine in question. GDSs are shown to be identical with problem solvers, controllers, and concepts. They may be composed to form more complicated GDSs, the result, to the first approximation, being an hierarchical structure (akin to the TOTE hierarchies of Miller *et al.* [35], in which lower-level (subgoal) GDSs are subroutines called by higher-level (goal) GDSs. This ordering yields a hierarchy of *abstraction* in the sense of Sections 1C and 1E. It is essential to recognize that it is *not* a control hierarchy with levels L^0, L^1. For a control hierarchy also exists and may, at any one level (L^0, L^1) have GDSs of arbitrary abstractional complexity.

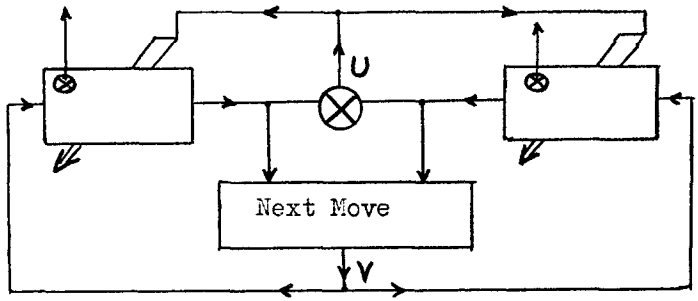

FIG. 7. A simple type of learning model. As in Fig. 5 the contents of the boxes are suppressed.

(The appropriate measuring procedure for sampling the levels of abstraction at *each* level of *control* is the combined sampling method described at the end of Section 1E.) Indeed, a major postulate of the theory is that at least two levels of control always do exist; GDSs do not exist in isolation; a GDS is invariably embedded in the domain of a higher-level GDS, which acts upon it and *rewrites* or *reproduces* it. So, for example, L^1 GDSs solve the higher-level problem of constructing L^0 GDSs that can achieve pertinent environmental goals.

A symbolic representation of the simplest type of learning and teaching is obtained by transforming Fig. 3 to yield the structure in Fig. 7. Information previously fed to the observer is now returned to α and β where it generates constructive operations. This information is a descriptive metastatement about the reaction of two individuals, α and β or learner and teacher. (Recall that metastatements are relativistic statements about individuals and that this is *why* they have the "meta" status. To comprehend the general case of conversational learning (Section 1G) the same

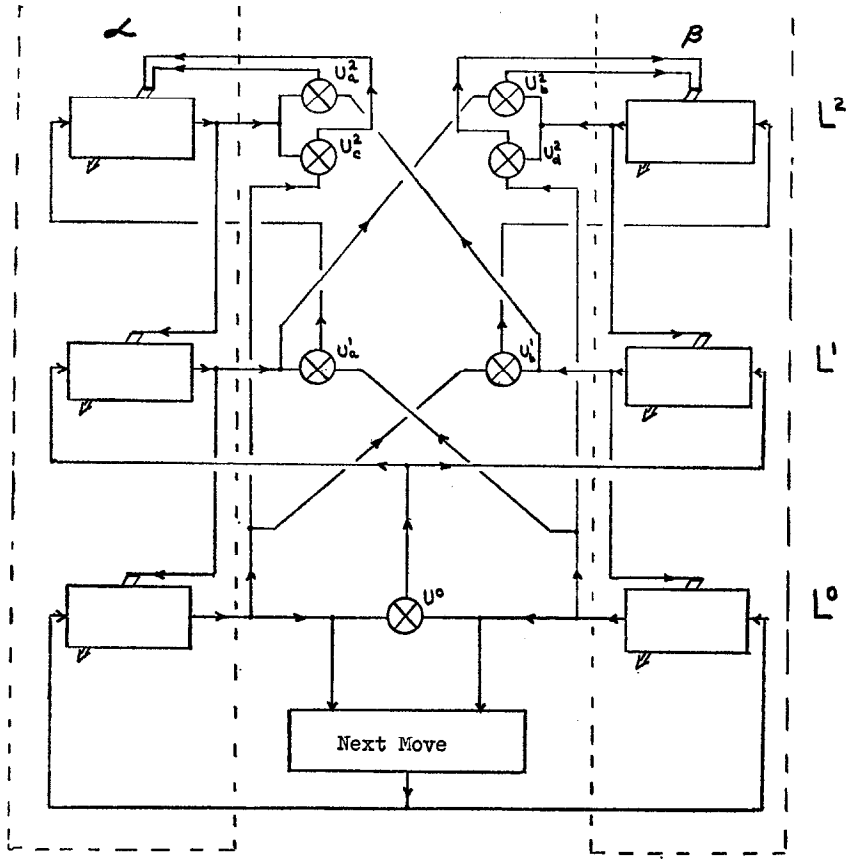

FIG. 8. The organization of multilevel α, β interaction. In particular α may be a student and β a teacher. The comparator outputs are labeled as in Fig. 5. It would complicate the figure to show further cross comparisons, such as those mentioned in Section 1E, but they generally exist. The contents of the boxes are suppressed for the same reason. U_a^2 and U_b^2 are fed directly back to the constructive operation at their own level because payoff invariance is assumed (as in Section 2E). The payoff invariance reflects the intuitive fact that (at this level) α is considering β's model of himself and vice versa, thus in a sense closing the system.

transformation is applied for each of the metalevels in Fig. 5, thus producing Fig. 8, which shows the organization used in the learning and teaching theories just examined.

In the case of "learning and teaching," the proper signals for regulating a reproductive or evolutionary process are available. But some difficulty is encountered in talking of individual and untutored learning (or, indeed, about individual mentation in general). To do so, it is necessary to coalesce α and β of Fig. 8 into a single entity; that is, to place α and β in a circumscribing box, to label the whole thing "the individual," and to equate this system with Fig. 6. The picture can be reasonably interpreted, of course. For example, at the lowest level of Fig. 6, U^0 is interpreted as an "internalized goal" (feeling I am right) signal obtained by confirming that two concurrently run problem-solving methods yield the same result and at the highest level $U_a^2 \wedge U_b^2$ is the "feeling of understanding a problem area." Similar interpretations are valid for more general situations. For example, those of Section 1: an individual's introspections or descriptive metastatements are made by comparing his past status (α) with his present status (β). The trouble is that in equating Fig. 8 with Fig. 6, we have changed the meaning of "individual." By hypothesis, Fig. 6 represents an individual. But, clearly, it is not α or β of Fig. 8, since it is, by definition, a shorthand representation for an organization containing both of them.

F. *Individuals*

Any discussion of interpersonal interaction hinges upon the concept of an individual, which, until the end of the last section, was taken for granted as a primitive construct. As is customary, α and β were interpreted as separate people; two heads, two brains, differently located in space. In the sense of Section 2E iv, α and β are separate "computing machines."

Now the customary distinction is perfectly sound. An individual (in any sense at all) is a stable and persistent entity. Clearly, α and β are stable and persistent, *because* their parts are physically replicated and homeostatically regulated. But this hardware-oriented method of demarcating individuals is not the *only* method and is not always the most propitious one. Indeed, in certain circumstances, it is *not*. Take the example of Section 2E v; if we are anxious to say "the (individual) student learns skill Z," then the "individual" referred to is not really α (a unique computing machine); the process called "learning Z" has constituents that are lodged in both α and β. Von Foerster points out that the teacher does much of the learning and means it (as I do) quite literally; someone who is said to "learn on his own" contains an "internal teacher." Similar dilemmas of *distribution* arise in discussing the ownership of the concepts that crop up during an α, β conversation if we wish to say "a concept Z belongs to him"

(an individual). In one sense any new concept is a product of the α, β conversation as a whole. In another sense, it belongs to a nameable entity; a *psychological* individual who is highly correlated with the computing machine α, but who is, in the course of the conversation, distributed between α and β.

Indeed, in many cognitive and social studies the psychological individual appears to be a more appropriate unit than the "piece of hardware unit" α or β. A corporation, a group, a culture, and the information structure named "Dickens" (the author), all have organizational integrity, though none of them inhabit a fixed brain or artifact. The psychological individual is thus a potentially useful unit of the caliber "distinct and nameable system." It remains to provide a hardware-independent definition of its boundaries.

G. *Psychological Individuals (P-Individuals)*

One definition of a P-individual emerges from the previous discussion. Since a P-individual is a persistent *organization*, rather than a persistent physical object, its boundaries will be set by a hardware-independent process that images the physical reproduction and regulation that maintain the integrity of the hardware individuals α and β. The requisite process has already been introduced; the reproduction of GDSs or, in general, of automata, and with this in mind we delineate a nesting of the systems that make up a P-individual.

The basic units of mentation have been characterized (in one theory of Section 2E iv) as GDSs, which may be concisely reinterpreted as "systems that preserve, maintain, or reproduce a *relation* (called the goal relation)." Similarly (Section 2E iv) a *memory* within the same theory is the reproduction of a GDS in the context of a higher-level GDS that acts as a constructive automaton. A recent paper by Loefgren [30] suggests a useful extension of this train of argument. Loefgren makes a distinction between "reproduction in context" and "self reproduction" (likewise, self explanation), pointing out that self reproduction may be formalized. Loefgren's paper is primarily addressed to refuting the Rosen paradox that a self-reproducing automaton, or a function that is a member of its own domain or range, is formally inconsistent. I do not attempt to outline Loefgren's argument, which is closely knit and elegant. (The interested reader is referred to the original document.) But I propose to define a P-individual as a collection of GDSs such that the entire collection is self-reproducing in the sense of Loefgren, thereby restricting the *possible* contents of the boxes in Fig. 6 (i.e., the GDSs they might contain).

In particular, a P-individual is generally a *symbiotically* self-reproducing system (one of several types noted by Loefgren) in which two sets of

organizations (the α-like entity and the β-like entity obtained when Fig. 8 is coalesced to produce Fig. 6) serve to reproduce one another. The internal interactions between the symbionts are precisely the self-referential statements and the constructive operations that were noted in Section 2E v as the proper regulators of reproduction and evolution.

H. *General Discussion (Formulations Similar to the P-Individuals)*

On inspection of Fig. 6 and Fig. 8, it is evident that the P-individual is really an *heterarchy* of control, rather than an hierarchy of control (though α will act like an hierarchy when it engages in interaction with another P-individual, β. As external observers, we *impose* a hierarchy of control in order to "freeze" our picture of things so as to avoid the difficulties that would occur if we allowed *any* GDS to reconstruct itself; the same difficulties that are encountered in talking about "impure programming." These difficulties also beset α, β communication and, consequently, the α, β language is stratified into referential levels, corresponding to levels of control, a point that has been made, quite generally, by Gorn [33]. In contrast, the communication between a pair of P-individuals takes place in an unstratified language, presumably in a natural language. Though generally unhampered, it must be restricted by devices for *local* or *contextual* stratification: see subsequent comments in the text.)

One consequence of this observation is that individuals no longer appear as goal-*seeking*, but goal-*setting* systems. From one point of view, the precondition for a P-individual is the existence of potentially concurrent processes that would be in conflict if actually executed together (i.e., at the same time and in the same place). This conflict reflects the fact that separate symbiotic systems must be distinguished if the control signals of a reproductive or evolutionary system are to be defined. A P-individual *is* a relationship between two or more processes that may be styled "the contemplation of" these processes. Its (his) bounds are set by the requirement of self-reproduction and its (his) integrity is maintained by the conflict in question. If the processes are run in different brains, then the P-individual is correlated with α and may see himself in relation to β. But this is a special case. In general, a P-individual is the *contemplation*, at the same time, of two or more simultaneously incompatible goals, unqualified. For example, "X attends to Y" is only meaningful if there is some other goal that X *might* aim to achieve, X *being* the contemplation of both boals or an indefinite number of them. Elsewhere [34] I have argued that the choice immanent in such an organization is the unique correlate of awareness, but this point will not be developed in the present article. It is more profitable to notice that a psychological individual is a primitive, stripped-down exemplar of one of Vickers' [36] Appreciative Systems.

To take a final tack, Fig. 6 is the cognitive analogue of the organization that Kilmer *et al.* [31] use to simulate the reticular orienting system of any vertebrate animal. In that context, the internal dialogue entailed by symbiotic self-reproduction is a manifestation of McCulloch's [32] "redundancy of potential command." It is interesting too the same structure embodied in the "proposer" and "critic" subprograms of several recent problem solving systems.

REFERENCES

1 G. Pask and B. N. Lewis, The self organisation of a three-person task-oriented group, *Proc. NATO Symp. Simulation of Human Behaviour* (Paris, 1967), *The simulation of human behaviour*, F. Bressow and M. de Mowtmallin, eds., pp. 291–311. Dunod, Paris, 1969.

2 G. Pask and B. N. Lewis, An adaptive automaton for teaching small groups, *Percept. Motor Skills* **14**(1962), 183–188.

3 G. Pask and B. N. Lewis, The development of communication skills under adaptively controlled conditions, *Programmed Learning* **2**(1964), 69–88.

4 E. B. Hunt, J. Marin, and P. J. Stone, *Experiments in induction*. Academic Press, New York, 1966.

5 D. Bannister and J. Mair, *The evaluation of personal constructs*. Academic Press, New York, 1968.

6 G. A. Kelly, *The psychology of personal constructs*, Vols. I and II. Norton, New York, 1955.

7 D. N. Hinkle, The change of personal constructs from the viewpoint of a theory of implications, Unpublished Ph.D. thesis, Ohio State University.

8 G. Bateson *et al.*, Towards a theory of schizophrenia, *Behavioural Sci.* **4**(1956), 251.

9 W. M. Brodey, *Changing the family*. Potter, 1968.

10 R. D. Laing, H. Phillipson, and A. R. Lee, *Interpersonal perception*. Tavistock, 1966.

11 G. Pask, Report 1: Road Research, Lab. Contract CON/1488/SU, June, 1970.

12 G. Pask, Fundamental aspects of educational technology (illustrated by the principles of conversational systems), *Proc. IFIP World Conf. Computer Education* (Amsterdam, 1970). (Vol. I, Invited Papers). R. Scheepmaker, ed. 129–152.

13 W. R. Ashby, *An introduction to cybernetics*. Chapman & Hall, London, 1956.

14 R. D. Luce and H. Raiffa, *Games and decisions*. Wiley, New York, 1957.

15 R. B. Banerji, *Theory of problem solving*. American Elsevier, New York, 1970.

16 N. Howard, The theory of metagames, *Gen. Systems Yearbook* **11**(1966), 167–186.

17 N. Howard, The mathematics of metagames, *Gen. Systems Yearbook* **11**(1966), 187–200.

18 A. Rappaport, Escape from Paradox, *Sci. Amer.* June (1966).

19 L. Festinger, *A theory of cognitive dissonance*. Tavistock, 1959.

20 H. von Foerster, What is memory that it may have hindsight and foresight, *BCL* 153 *Tech. Rep.* (1968).

21 H. von Foerster, Molecular ethology, *Tech. Rep. BCL*. To be published in *Molecular mechanism of memory and learning* (C. Ungar, ed.). Plenum Press, New York, 1971.

22 G. Pask, A cybernetic model for some types of learning and mentation, in *Cybernetic problems in bionics* (H. C. Oestreicher and D. R. Moore, eds.), pp. 531–585. Gordon & Breach, New York, 1968.

23 G. Pask, Adaptive Machines, *Proc. NATO Symp. Major Trends in Programmed Learning Res.* (Nice 1968). *Programmed Learning Research*, Dunod, 1969, pp. 251–261. (F. Bressow and M. de Mowtmallin, eds.).
24 J. von Neuman, *Theory of self-reproducing automata.* Univ. of Illinois Press, 1966.
25 G. F. Brieske, Analysis and synthesis of a learning teaching system, *BCL Tech. Rept.* **16**(1969), Vol. 1.
26 L. J. Fogel, A. J. Owens, and M. J. Walsh, *Artificial intelligence through simulated evolution.* Wiley, New York, 1966.
27 A. Newell, A variety of intelligent learning in a general problem solver, in *Self organising systems* (M. Yovitts and S. Cameron, eds.), pp. 153–186. Macmillan (Pergamon), New York, 1960.
28 G. Pask, The computer simulated development of populations of automata, *Math. Biosci.* **4**(1969), 101–127.
29 N. Rescher, *The logic of commands.* Routledge and Keegan Paul, 1967.
30 L. Loefgren, An axiomatic explanation of complete self reproduction, *Bull. Math. Biophys.* **30**(1968), 415–425.
31 W. L. Kilmer, W. S. McCulloch, *et al.*, On a cybernetic theory of the reticular formation, in *Cybernetic Problems in Bionics* (H. L. Oestreicher and D. R. Moore, eds.), pp. 431–479. Gordon & Breach, New York, 1968.
32 W. S. McCulloch, *Embodiments of mind.* MIT Press, Cambridge, Massachusetts, 1965.
33 S. Gorn, The treatment of ambiguity and paradox in mechanical languages, in *Recursive function theory* (Symposia in Pure Mathematics, Vol. 5). Amer. Math. Soc., Providence, R.I., 1962.
34 G. Pask, The cybernetics of behaviour and cognition, Extending the meaning o "goal" in *Progress in Cybernetics*, Vol. I, pp. 15–44. Gordon & Breach, New York 1970.
35 G. A. Miller, E. Gallanter, and K. Pribram, *Plans and the structure of behaviour.* Holt, Rinehart & Winston, Dryden, 1960.
36 G. Vickers, *Value systems and social process.* Pelican, 1970.